Amendments to the Specification:

Please replace the paragraph beginning at page 2, line 29 to page 3, line 3 with the following amended paragraph:

Although the above methods exist, there is a need for methods for rapid and economical testing of large molecular libraries so that better candidate drug molecules can be discovered. Also there remains a need for new methods and tools to design linear epitopes that can be specifically and tightly bound by capture agents. Therefore, among the objects herein, it is an object to provide such methods and products.

Please replace the paragraph beginning at page 4, lines 5-10 with the following amended paragraph:

In another embodiment, the collection of polypeptides ~~contain~~ contains at least two polypeptides that contain the same four critical residues but each of the two polypeptides has non-critical residues at different positions. The polypeptides contain at least 6 unique residues and at least 2 non-critical residues that are adjacent to each other. In one example, the non-critical residues are selected from among Y, S and G.

Please replace the paragraph beginning at page 5, lines 11-21 with the following amended paragraph:

Also provided herein are collections of capture agent - binding partner polypeptide pairs containing a collection of polypeptides that [[can]] are antigenic, such as described herein for use as binding partners and a collection of capture agents. Each capture agent in the ~~collectino~~ collection binds to a binding partner polypeptide within the collection of binding partner polypeptides. In one example, the capture agents of the collection are antibodies or antibody fragments. The collection of binding partner polypeptides can include 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 50, 100 or more of polypeptides of any of SEQ ID Nos. 1-911. The collections typically

include at least 10 members, including at least two of the polypeptides of any of SEQ ID Nos. 1-911.

Please replace the paragraph beginning at page 6, line 21 to page 7, line 4 with the following amended paragraph:

The methods provided herein further include generating highly ~~antigenic~~ antigenic, highly specific polypeptides containing critical and non critical amino acids. In one embodiment, the methods include generating a subset of polypeptides of length q residues, wherein q = m + r and r is the number of non-critical amino acids, wherein r is an integer equal to or greater than 1 and q is an integer greater than 4. In another embodiment, the N and C terminal amino acids of the polypeptides of length q residue are critical amino acids. In yet another embodiment at least 2 of the non-critical amino acids are adjacent in the polypeptide. The methods include generating polypeptides with any number of non-critical amino acids. For example, the number of non-critical amino acids r can be 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10. The methods include generating highly antigenic highly specific polypeptides of any length. [[for]] For example, the length of such polypeptides ~~include~~ includes polypeptides of length q, where q is an integer between 5 and 100 or 5 and 50 or 5 and 30 or 5 and 20 or 5 and 10.

Please replace the paragraph beginning at page 7, lines 5-18 with the following amended paragraph:

The methods for generating highly ~~antigenic~~ antigenic, highly specific polypeptides provided herein also include selecting a subset of dissimilar polypeptides. For example, dissimilarity refers to functional and structural dissimilarity based upon predetermined criteria. Dissimilarity is assessed by comparing each polypeptide in the set S1 an arbitrarily selected reference polypeptide from the set S1 by comparing corresponding critical ~~residue~~ residues based upon position in the polypeptides. Polypeptides from set S1 are selected that contain residues most dissimilar from the reference polypeptide. In one embodiment, dissimilarity is determined by calculating a similarity score from a similarity matrix by comparing values for the

corresponding critical residues in the reference polypeptide to the corresponding critical residues in the polypeptides of set S1, combining the scores for the residues in each polypeptide to generate a score for each polypeptide and selecting those below a predetermined score.

Please replace the paragraph beginning at page 7, lines 19-29 with the following amended paragraph:

Also provided herein are collections of binding partner polypeptides, comprising 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 50, 100 or more polypeptides generated by methods of producing highly antigenic highly specific polypeptides. Also provided herein are collections of capture agent - binding partner polypeptide pairs comprising collections of binding partner polypeptides generated by the methods herein and collections of capture agents. In such collections, the capture agents each bind to a binding partner polypeptide within the collection of binding partner polypeptides. The collections can be contained in kits that optionally ~~including~~ include instructions for preparing capture agents that specifically bind to members of the collection.

Please replace the paragraph beginning at page 9, line 9-16 with the following amended paragraph:

The methods include optionally synthesizing polypeptides of length d, containing n fixed amino [[acids]] acid positions, where the polypeptides of the collection have the same amino acid at a fixed position. The number of fixed positions can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12. The number of variable positions included for synthesis can be 2, 3, 4, 5, 6, 7 or 8. In one embodiment, the synthesized polypeptides are highly antigenic highly specific polypeptides, such as sequences of highly antigenic highly specific polypeptides generated by the methods provided herein.

Please replace the paragraph beginning at page 9, line 17 to page 10, line 8 with the following amended paragraph:

Also provided herein are methods for synthesizing an addressable collection of molecules. The methods include providing a collection of b tags and a collection of b addressable capture agents, where each capture agent binds a unique tag and b is the number of tag-capture agent pairs. The tags are presented in an addressable format suitable for chemical synthesis. A collection of molecules is synthesized such as on starting molecules and each tag is conjugated directly or indirectly via a linker to a starting molecule. Each synthesized molecule contains a number of variable constituent positions X conjugated to the starting molecule. The method of synthesis includes synthesizing a subset of X positions X_1 in a first round of synthesis to generate a collection of tag- X_1 molecules, whereby each unique tag is conjugated to a unique combination of constituents at the synthesized X_1 positions. The collection of synthesized tag- X_1 molecules is mixed and split into b addressable first-round ~~subsets,such~~ subsets, such that each first-round subset contains a collection of tag- X_1 molecules representing on average every possible combination of constituents at the synthesized X_1 positions. A further subset of constituent positions X_2 is synthesized in a further round of synthesis, such that each first-round subset is conjugated to a unique combination of constituents at X_2 positions to generate b second-round subsets. The resulting tag-X_1X_2 is contacted with an addressable collection of b capture agents to produce an addressable collection of synthesized molecules. In one example, the synthesized molecules are selected from among nucleic acid molecules, polymers, biopolymers, polypeptides, and small organic molecules. In one example, the starting molecule is a pharmacophore. In another example, the starting molecule is a monomer and the synthesized molecules are polymers. In another example, the tags are highly antigenic highly specific polypeptides. In one example, the tags comprise any of the sequences set forth in SEQ ID NOs. 1-911. In yet another example, the capture agents are antibodies or antibody fragments.

Please replace the paragraph beginning at page 12, lines 9-20 with the following amended paragraph:

As used herein, [[an]] a highly antigenic, highly specific polypeptide (also referred to herein as HAHS polypeptides) is a polypeptide that specifically binds to a unique member of a

collection of capture agents (i.e. binds with at least 1-, 2-, [[5-]] 5-, 10-fold or greater affinity to one unique member compared to all other members in a collection of at least 3, 5, 10, 50, 100 or more unique members). Collections of HAHS polypeptides are collections of polypeptides that specifically bind capture agents such that in collections thereof each HAHS ~~polyeptide~~ polypeptide in the collection will bind to a unique member of a collection of capture agents with greater affinity (~~tyically~~ typically at least 1, 2, 5, 10-fold or more) than to any other member of the collection of capture agents. The collections of capture agents include at least 3, 5, 10, 50, 100 or more unique capture agents.

Please replace the paragraph beginning at page 13, lines 8-13 with the following amended paragraph:

As used herein, antigenic when used in the context of highly ~~antigenic~~ antigenic, highly specific polypeptides refers to polypeptides that induce, upon administration to a host, antibodies that are specific for the HAHS polypeptides or upon screening, or select for (such in display or panning methods) capture agents, such as antibodies or antibody fragments, with specific and selective binding to the HAHS polypeptides.

Please replace the paragraph beginning at page 13, lines 14-25 with the following amended paragraph:

As used herein, a molecule, such as capture agent, that specifically binds to a polypeptide, such as a HAHS polypeptide provided herein, typically has a binding affinity (K_a) of at least about 10^6 l/mol, 10^7 l/mol, 10^8 l/mol, 10^9 l/mol, 10^{10} l/mol or greater (generally 10^8 or greater) and binds generally with greater affinity (typically at least 10-fold, generally 100-fold or) than to the molecules and biological particles that are to be detected or assessed in the methods that employ the capture systems. Thus, affinity refers to the strength of interaction between two or more molecules, such as a capture agent and a HAHS polypeptide binding partner. Typically, an HAHS polypeptide specifically binds to a unique capture agent in

collection with at least 1-, 2-, [[5-]] 5-, 10-fold or greater affinity than to all others capture agents in a collection.

Please replace the paragraph beginning at page 18, lines 23-38 with the following amended paragraph:

As used herein, the term "polypeptide" is used interchangeably with the term "protein" and includes peptides containing two or more amino acids. A polypeptide can be a single polypeptide chain, or [[to]] two or more polypeptide chains that are held together by non-covalent forces, by disulfide cross-links, or by other linkers (*e.g.* peptide linkers). Thus, a single heavy or light chain of an antibody, or an antibody fragment containing all or part of the heavy and light chains of an antibody, no matter how the chains are associated or joined, are exemplary molecules that are included within the term "a polypeptide." A polypeptide can contain non-proteinaceous components, such as sugars, lipids, detectable labels or therapeutic moieties. A polypeptide can be derivatized by chemical or enzymatic modifications (*e.g.* by replacement of hydrogen by an alkyl, acyl, or amino group; esterification of a carboxyl group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative; phosphorylation or dephosphorylation of a serine, threonine or tyrosine residue; or N- or O-linked glycosylation) or can contain substitutions of an L-configuration amino acid with a D-configuration counterpart.

Please replace the paragraph beginning at page 23, lines 14-17 with the following amended paragraph:

As used herein, normalization refers to the equilibration of the titer or concentration of all members of a library, such as a tagged library, so that the number of particular members, such as tagged members or total members, in two samples or portions are about the same.

Please replace the paragraph beginning at page 33, line 30 to page 34, line 9 with the following amended paragraph:

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors." In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. "Plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. Other such [[other]] forms of expression vectors that serve equivalent functions and that become known in the art subsequently hereto.

Please replace the paragraph beginning at page 37, lines 4-10 with the following amended paragraph:

Methods also are provided herein for generating or selecting capture agents which bind to highly ~~antigenic~~ antigenic, highly specific polypeptides. Such methods include introducing collections of HAHS polypeptides into an animal and isolating antibodies as a result of raising an immune response to the introduced HAHS polypeptides. The methods also include selecting capture agents from a collection of candidates for the capture agents which selectively and specifically bind to one or more HAHS polypeptides.

Please replace the paragraph beginning at page 37, lines 12-22 with the following amended paragraph:

Provided herein are methods for obtaining highly ~~antigenic~~ antigenic, highly specific (HAHS) polypeptides for use as partners with capture agents such as antibodies. The polypeptides contain any number of amino acids against which a specific capture agent can be generated, selected or synthesized to bind. Typically such polypeptides are at least 2, 3, 4, 5, 6, 7, 8 to about 100 amino acids in length, usually between 2-50, 2-40, 2-30, 2-20, 4-20, 5-20, 2-

50, 4-50, 5-50, and 6-20 amino acids in length. Also provided are methods for generating capture agents, such as antibodies, which bind to HAHS polypeptides. Thus, methods generate pairs of HAHS polypeptides and capture agents. There is no detectable cross-reactivity, such as by ELISA assay, between or among different pairs of HAHS polypeptides and capture agents.

Please replace the paragraph beginning at page 41, lines 4-27 with the following amended paragraph:

Antigenic amino acids can include natural and/or non-natural amino acids, such as non-natural amino acids described further herein. Non-naturally occurring amino acids can be ranked for antigenicity using methods applied to the naturally occurring amino acids, for example by testing sequences against antisera or libraries of antibodies (described herein) and can be ranked along-side naturally occurring amino acids. For example, a representative set of polypeptides composed of non-naturally occurring amino acids and/or a combination of non-naturally occurring and naturally occurring amino acids of a chosen polypeptide length can be used to immunize animals. Based on the subset of polypeptides injected which are antigenic and non-antigenic, amino acids are identified which either are more likely to be present in antigenic polypeptides or are more likely to be present on non-antigenic polypeptides. The likelihood of [[a]] an amino acid's presence in antigenic polypeptide gives an observed antigenic ranking. Some non-natural amino acids are very structurally similar to naturally occurring amino acids and to other non-naturally occurring amino acids. This similarity can be factored in to provide antigenicity rankings based on these similarities. For example, a collection of polypeptides can be generated containing non-natural amino acids and tested for antigenicity. Polypeptides which are antigenic can be used to create further sets of polypeptides (replacement sets) by systematically replacing some or all of the amino acids systematically to determine which amino acids are critical. The data can then be analyzed for the replacement sets to determine a factor for each non-natural amino acid, where the factor represents the frequency of finding the particular non-natural amino acid in a critical position within an antigenic polypeptide.

Please replace the paragraph beginning at page 41, line 28 to page 42, line 2 with the following amended paragraph:

The use of non-naturally occurring amino acids increases the diversity and thus uniqueness of the polypeptides that can be generated. For example, there are several hundred non-naturally occurring amino acids that are commercially available and [[a]] an even larger number that can be synthesized by standard chemistry methods known in the art. Non-naturally occurring amino acids can be used at either critical or non-critical residues or at both critical and non-critical residues. The ability to incorporate non-naturally occurring amino acids also permits linear, cyclic and branched polypeptide structures to be designed and constructed.

Please replace the paragraph beginning at page 42, lines 20-26 with the following amended paragraph:

Non-natural amino acids that are modifications of natural amino acids in the side chain functionality, such that the methylene groups of the side chain of the natural amino acids have been substituted by imino groups or divalent non-carbon atoms or, alternatively, methyl groups have been substituted by amino groups, hydroxyl groups or thiol groups, so as to add ability to form hydrogen bonds or to reduce their hydrophobic properties (*e.g.* leucine to 2-aminoethylcysteine, or ~~isolecine~~ isoleucine to o-methylthreonine).

Please replace the paragraph beginning at page 51, line 1-7 with the following amended paragraph:

Methods for preparing collections of HAHS polypeptides in [[a]] an addressable format

Provided herein are methods for preparing collections of HAHS polypeptides in an addressable format. The methods are flexible for collection size and include preparation of small and large polypeptide collections, including large diverse collections of HAHS polypeptides, addressably formatted and displayed suitable for screening and other assays.

Please replace the paragraph beginning at page 54, lines 14-20 with the following amended paragraph:

The second round of synthesis results in tagged AB peptides further extended with a unique combination at positions C and D, such that each AB combination has been extended with each CD possibility. A total of b x b ~~polypeptides~~ polypeptide possibilities have been synthesized. The SB positions are identifiable by the tags, since each AB possibility is linked to a unique tag. The CD positions are identifiable by their position in the second round synthesis, each address represents a unique CD combination.

Please replace the paragraph beginning at page 55, lines 22-30 with the following amended paragraph:

Pairs of tags and capture agent pairs are assembled and conjugated to beads. The number of pairs is chosen to be 100 (b = 100). The 100 tags are conjugated to a solid support, such as latex beads, and distributed to the wells of a plate, gridded so that they are arranged in a predetermined 10 x 10 or other suitable format, predetermined so that it is known which tag is at which position in the grid. The tags can optionally be conjugated to a linker such as GS or [[GSG.,]] GSG, such that the synthesized peptide is represented tag-GS-ANNBCD or tag-GSG-ANNBCD, respectively. Standard solid-phase peptide synthesis chemistry is employed to synthesize the tagged peptides.

Please replace the paragraph beginning at page 56, lines 10-26 with the following amended paragraph:

The mix is redistributed to a second synthesis block such that essentially all of the combinations from the first block represented at each synthesis address (*e.g.* each well, tube etc) in the second block. The second block also is gridded out as a 10 x 10 or other suitable format such that 10 amino acids will be distributed at each of the third and fourth variable positions. Position C is synthesized by adding 10 amino acids to the synthesis grid as follows. Each row receives 1 amino acid, such that all positions in the row receive the same amino acid

and each different row receives a different amino acid[[.]] (*e.g.* row 1 receives amino acid 1, row 2 receives amino acid 2 etc). Position D is synthesized by adding the 10 amino acids designated for position D to the grid where each column receives 1 amino acid, such that all positions in the column receive the same amino acid and each different column receives a different amino acid[[.]] (*e.g.* column 1 receives amino acid 1, column 2 receives amino acid 2 etc). The peptides synthesized are represented by the formula tag-linker-A-GS-B-C-D, where addresses on the grid can be represented as the series: tag-linker-A_{1-10}-GS-B_{1-10}-C_1-D_1, tag-linker- A_{1-10}-GS-B_{1-10}-C_1-D_2....A_{1-10}-GS-B_{1-10}-C_2-Dy...A_{1-10}-GS-B_{1-10}-C_{10}-D_{10} [[(sees]] (see Figures 3A and 3B).

Please replace the paragraph beginning at page 62, lines 3-14 with the following amended paragraph:

Capture agents also can be constructed from complementarity determining regions (CDRs). Recombinant means can be used to isolate the CDRs which are contained in the hyper variable loops of the antibody variable domain and are involved in antigen binding. Once isolated one or (up to all 6) of the CDRs can be cloned into a protein scaffold (see for example, Skerra (2000) *J. Mol. Recognit.* 13:167-187). Protein scaffolds include any polypeptide in which the CDRs can be placed and maintain binding to an antigen. Exemplary protein scaffolds include antibody and antibody fragments, fibronectin, protease inhibitors such as bovine pancreatic trypsin inhibitor, human pancreatic trypsin inhibitor, and ~~tendmsitat,~~ tendamistat, helix bundle proteins including natural and engineered structures such as the "Z" domain, lipocalins, knottins, and enzymes such as glutathione S-transferase, thioredoxin, and triose phosphate isomerase.

Please replace the paragraph beginning at page 65, lines 17-30 with the following amended paragraph:

Cross-linking reactions involving molecules and binding partners, such as proteins, are generally reactive group reactions, such as side chain reactions, and are nucleophilic, resulting in

a portion of the end of the cross-linker being displaced in the reaction (the leaving group). Nucleophilic attack is dependent on the pH, temperature and ionic strength of the cross-linking buffer. For example, when the buffer is one to two pH units below the pK_a of the reactive group, such as a side chain, the species is highly protonated and is most reactive. One to two [[Ph]] pH units above the pK_a, the species is not protonated and not reactive. The majority of molecules and binding partners, such as proteins, have reactive groups, such as primary amines and free ~~sulfhydrals,~~ sulfhydryls, available at the surface or terminus of the molecules or binding partner. These are the two most commonly used groups in molecular cross-linking strategies. Cross-linking strategies also can use carbohydrates, carboxyls or other reactive functional groups.

Please replace the paragraph beginning at page 66, lines 12-24 with the following amended paragraph:

Fusion proteins can be produced by recombinant expression of nucleic acids that encode the fusion protein. The formation of a fusion protein involves the placement of two separate coding sequences, such as genes or ~~nucleotides~~ nucleotide sequences, one encoding the displayed molecule and the second encoding the binding partner, in sequential order in an appropriate cloning vector. Methods for creating an expression vector containing the displayed molecule and the binding partner are well known to those of skill in the art (see, *e.g.*, Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, Clod Spring Harbor Laboratories, Cold Spring Harbor, New York). Additional methods for the formation of a fusion protein conjugate include, but are not limited [[to]] to, ligation of sequences resulting in linear tagged cDNA molecules; primer extension and PCR for binding partner incorporation; insertion by gene shuffling; recombination strategies; incorporation by transposases; and incorporation by splicing.

Please replace the paragraph beginning at page 70, lines 28-31 with the following amended paragraph:

As described above, HAHS polypeptides [[cam]] can be used as binding partners to tag biological molecules and particles. The methods provided herein can be used to generate

collections of binding partners to which capture agents, such as antibodies and antibody fragments bind.

Please replace the paragraph beginning at page 71, lines 18-25 with the following amended paragraph:

Nucleic acid encoding a HAHS polypeptide binding partner also can include sequences of nucleotides that can aid in unique or convenient priming, such as for PCR amplification, or can encode amino acids that confer desired properties, such as trafficking signals, detection, solubility alteration, facilitation of purification or conjugation or other functions or provide other functions. For example, in embodiments in which candidate components are subcloned into a panel of vectors each containing an HAHS binding partner, these additional sequences also can be included in the vector.

Please replace the paragraph beginning at page 74, lines 19-25 with the following amended paragraph:

Numerous methods have been developed for the immobilization of proteins and other biomolecules onto solid or liquid supports. Among the most commonly used methods are absorption and adsorption or covalent binding to the support, either directly or via a linker, such as the numerous disulfide linkages, thioether bonds, hindered disulfide bonds, and covalent bonds between free reactive groups, such as amine and thiol groups, known to those of skill in the art.

Please replace the paragraph beginning at page 75, lines 23-29 with the following amended paragraph:

Collections of molecules and/or biological particles can be screened using HAHS polypeptides in capture systems, such as described herein, or in any other screening means ~~know~~ known in the art. In preparation for screening, collections of molecules and/or biological particles can be generated and tagged with HAHS polypeptides. Such tagged molecules and/or

particles can be displayed for example on a solid support, for example, through interactions with capture agents. The collections can then be screened for functions or effects of interest.

Please replace the paragraph beginning at page 78, lines 9-17 with the following amended paragraph:

For each variable position, a set of substituents is chosen, each set represents all of the possible substituents to be added at that position. The number of substituents chosen for use in synthesizing the two positions is set by the total number of available tag:capture agent pairs, b, ~~such that~~ such that $b = X_A \times X_B$ where X_A and X_B are the number of substituents in the set of substituents for each of the variable positions to be synthesized in the first round. The first round of synthesis generates collections of molecules where each unique tag now has a unique ~~contains a unique~~ combination of substituents at the two variable positions.

Please replace the paragraph beginning at page 78, lines 18-26 with the following amended paragraph:

A second round of synthesis is initiated by mixing all the tagged synthesized molecules of the first round together and distributing them to a grid or otherwise divided synthesis container of b positions. The second round of synthesis adds an additional two variable positions of substituents, and optionally an additional number of fixed positions. The number of substituents chosen for use in synthesizing the two positions is set by the total number of available tag:capture agent pairs, b, such that ~~such that~~ $b = X_C \times X_D$ where X_C and X_D are the number of substituents in the set of substituents for each of the variable positions to be synthesized in the first round.

Please replace the paragraph beginning at page 79, lines 9-20 with the following amended paragraph:

Molecules synthesized at each address in the second round synthesis are incubated with a separate collection of addressed capture agents, such that there are b collections of addressed

capture agents, each containing the same capture agents. For example, a canvas of b capture agent arrays is used where molecules from each address at the second round are incubated with a separate array on the canvas. Such distributions generate collections of capture agents, each collection displaying a subset of the synthesized molecules and together displaying the full set of synthesized molecules. Each collection of capture agents displays molecules with a unique combination of substituents added in the second synthesis round and the full assortment of possibilities of substituents added in the first synthesis round. The displayed collections of synthesized molecules can be used for screening and other functional assays.

Please replace the paragraph beginning at page 80, lines 28-33 with the following amended paragraph:

The operations described above to generate collections of polypeptide sequences can be performed with the assistance of one or more computer programs (software) executing on a computer. The following description of a suitable computer system and software is [[an]] exemplary, for purposes of illustration only. Other suitable computer systems and software can be used by one of skill in the art to perform the methods.

Please replace the paragraph beginning at page 84, lines 5-15 with the following amended paragraph:

The operational process illustrated by the flow diagram of FIGURE 2 can be performed on the computer system illustrated in FIGURE 1 by using one or more computer ~~program~~ programs to run different software routines. It should be understood that all routines can be integrated into a single computer program or can be performed by multiple programs with [[and]] an arrangement of program steps. Programs to be employed rely on a suitable database of amino acid data, such as antigenicity and similarity rankings, from which amino acids are selected and from which amino acids and polypeptide sequences are compared. Such databases are readily available and those skilled in the art will be knowledgeable with regard to extracting the appropriate data (see for example, Geysen *et al.*, (1988). *J. Molecular Recognition* 1:32-41).

Please replace the paragraph beginning at page 84, line 17 to page 85, line 2 with the following amended paragraph:

The methods provided herein generate collections of HAHS polypeptides which can be utilized as a diverse collection of epitopes for diagnostic assays, such as diagnostics for diseases and conditions. For example, a collection of HAHS polypeptides is generated and used to assess the antibodies present in a sample, such as from an animal, subject or patient. Collection of HAHS polypeptides are generated in an addressable format, such as arrayed [[such]] on a solid support or associated with color-coded or tagged beads. The addressable collection of HAHS polypeptides is then contacted with samples containing antibodies. Samples can include any fluids, tissues and/or cells which contain antibodies and/or fragments of antibodies, such as but not limited to, blood, sera, spleen, lymph tissue, bone marrow, lymphocytes, plasma cells and B cells. Diagnostic assays can include assessing the number or pattern of HAHS polypeptides ~~bounds~~ and/or the amount of each HAHS polypeptide bound. Results can be compared between samples, or between a sample and a control. For example, a sample from a diseased subject can be compared with a control non-diseased sample. Subjects can be treated with an agent, such as a small molecule, a pathogen and or one or more antigens, samples collected and tested against the collection of HAHS polypeptides. Such samples can be compared with untreated controls to assess differences in antibody levels or types between treated and untreated samples.

Please replace the paragraph beginning at page 86, lines 25-34 with the following amended paragraph:

Step 4: Since 4 residues [[ere]] <u>were</u> selected from the total selected length of 6 (step 3), the remaining 2 residues, designated "non-critical" were then assigned. For exemplary purposes, the 2 non-critical residues were assigned adjacent positions and only critical residues were chosen to occupy the N-terminal and C-terminal positions, thereby generating the possible 6-mers into which non-critical residues were placed. For exemplification, two possible combinations of non-critical residues were selected. These were Tyr-Gly, and Ser-Gly. These

were chosen because they confer improved solubility and permit hairpin folding which can be advantageous for generating capture agents/binding partners for the methods and products herein.